Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 16, no. 7, July 2020, pp. 1078-1086
DOI: 10.23940/ijpe.20.07.p10.10781086

Using Genetic Algorithm to Augment Test Data for Penalty Prediction

Chunyan Xia[a,b], Xingya Wang[c,*], Yan Zhang[a], and Hao Yang[a]

[a]*Department of Computer and Information Technology, Mudanjiang Normal University, Mudanjiang, 157011, China*
[b]*State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, 210023, China*
[c]*College of Computer Science and Technology, Nanjing Tech University, Nanjing, 210009, China*

Abstract

With the development of smart court construction, a deep learning method has been introduced into the field of penalty prediction based on judicial text. Since the increasing parameters of the penalty prediction model, the size of the data set to test the performance of the model is gradually expanding. First, we use the data augmentation method to make some changes to the original data to obtain a large number of augmented data with the same label. Then, we use the multi-objective genetic algorithm to search for high-quality test data from a large number of augmented data, so as to improve the diversity of augmented data. Finally, we perform experiments. The results of actual judicial cases show that compared with the random method, augmented test data based on the genetic algorithm can better test the performance of the penalty prediction model.

Keywords: penalty prediction; data augmentation; test data; genetic algorithm

(Submitted on May 2, 2020; Revised on May 30, 2020; Accepted on June 29, 2020)

1. Introduction

With the advancement of the construction of court information 3.0, the judicial big data that can be stored and processed by the computer is growing rapidly. Judicial organs have introduced artificial intelligence into the field of legal services to realize the intelligence of judicial business [1-2]. Since Becker, most of the economic analysis has adopted the legal standard, taking the penalty as the main means of enforcement [3]. Practice also shows that people generally believe that the degree of punishment is the main factor for most parties to avoid hurting others [4]. Among all kinds of judicial data, judicial documents are the most common text information in judicial cases [5]. In recent years, some achievements have been made in the research of penalty prediction models in the judicial field, such as penalty prediction [4], automatic sentencing prediction [6], and the prediction of laws and charges [7], etc.

With the rapid development of machine learning technology, more and more attention has been paid to the combination of justice and deep learning. In order to improve the performance of the penalty prediction model, researchers continue to increase parameters of the training model, leading to a gradual expansion of the training set. Therefore, in order to improve the generalization ability of the prediction model, it is necessary to increase the number and diversity of test data in the test set. Using the data augmentation method, that is, on the basis of not actually increasing the original data, some small changes on original data can obtain more data, which can increase the number of test data with the same labels for the penalty prediction model, so as to improve the accuracy of the model prediction [7]. Traditional data augmentation is usually applied to the image field, which can increase the number and diversity of images by clipping, rotation and scaling. Generally, these methods are not suitable for the text data scene [8]. Because the word order in text data is very important information, the change of word order will lead to the change of the whole sentence meaning. Therefore, this paper takes sentences as the basic unit and augments the text data of describing cases by randomly changing the order of sentences in the text, randomly deleting sentences in the text and randomly adding sentences in other texts.

With the development of text analysis technology and the constant emergence of data augmentation methods, the scale

* Corresponding author.
E-mail address: xingyawang@outlook.com

of test set for judicial augmentation is gradually increasing. It is necessary to rely on search technology to select high-quality test data that can effectively test the performance of the model. As a type of heuristic search algorithm, the genetic algorithm (GA) is an adaptive search algorithm of global optimization probability formed by simulating the genetic and evolutionary process of biology in the natural environment. Its main goal is to find the optimal solution or approximate optimal solution of the problem in the search space quickly. It has been successfully applied in the field of multi-objective optimization [9]. In this case, GA is used to search the test data that meets the multi-objective optimization problem from the augmented large amount of data, so as to provide high-quality test data for testing the performance of the penalty prediction model.

In this paper, a method for augmenting penalty test data based on GA is proposed. First, on the basis of the original test set of the penalty prediction model, a large number of augmented data are obtained by scrambling, inserting, and deleting sentences in the text. Then, a multi-objective genetic algorithm is used to search for high-quality test data from a large number of augmented data, so as to improve the diversity of augmented test data. Finally, experimental results on the actual judicial case data show that the augmented test data based on GA can better test the performance of the penalty prediction model compared with the random method.

2. Technical Background

2.1. Penalty Prediction

Judicial data is related to people's livelihood, and judicial documents are the main research objects in the judicial field. With the rapid development of machine learning technology, a deep learning method is also introduced into the field of penalty prediction based on judicial text.

Most of the previous studies on the prediction of penalty regard it as a classified task. On the one hand, it aims to extract effective features and make use of them. Wang et al. [10] used the text classification method of in-depth learning to predict the term of penalty according to the actual case described in the judicial text. Liu et al. [11] classified the judgment documents by using the binary tree to establish the vocabulary. On the other hand, the successful method based on the neural network is used to build the model. Luo et al [12] proposed a neural network method based on attention to build a penalty prediction model. Zhong et al. [13] formalized the dependency relationship between judicial trial businesses into a directed acyclic graph and proposed a topological multi-task learning framework.

The text classification method can easily obtain a large number of label data. Therefore, compared with traditional methods, the deep learning method has achieved better results in the field of penalty prediction.

2.2. Data Augmentation

Data augmentation is a widely used technology in many machine learning tasks, which can virtually expand the size of the training data set and avoid overfitting [14]. In the image field, Ryo et al. [15] proposed an image preprocessing method based on random clipping and patching, which randomly clipped four images and patched them to create a new training image. Zhong et al. [16] introduced a new data augmentation method for training the convolutional neural networks. This method randomly selects the rectangular area in the image and erases its pixels with random values to generate training images with various occlusion levels, so as to reduce the risk of overfitting.

Natural language is composed of words, so it is difficult to obtain general conversion rules to ensure the quality of augmented data, and it can be applied to various fields. The existing methods of text data augmentation include translation transformation, synonym substitution, and text editing. Translation transformation refers to the use of translation tools to translate sentences into other languages, and then translate back to the original language to generate new data [17]. Synonym substitution refers to the random selection of some words in the text, replacing them with synonyms of high similarity [18]. Text editing refers to that part of the text content is repeatedly mentioned in a specific scene, and the edited text has the same meaning as the original text [7].

In the case of less available data, the method of data augmentation can effectively use existing data to increase the size of labeled data sets, so as to improve the performance of the deep learning model [19].

2.3. Multi-Objective Genetic Algorithm

The multi-objective genetic algorithm is an evolutionary algorithm used to analyze and solve the multi-objective optimization problem. Its core is to coordinate the relationship between the various objective functions and find out the optimal solution set that makes each objective function reach as large as possible. Among the multi-objective genetic algorithms, NSGA-II is a classical multi-objective optimization algorithm with the greatest influence and the widest application range [20].

There are few kinds of research on deep learning of GA. Ma et al. [21] generated a population for convolutional neural networks structure by selection, crossover, and mutation operations of GA to achieve the task of image classification and achieved good results.

It is necessary to combine the multi-objective genetic algorithm to search for high-quality test data, so as to better test the performance of the prediction model.

3. Method Framework

In this paper, the original data used to test the performance of the penalty prediction model is the structured label data of the judgment document. The case information is as follows: Charge, Date, Address, Criminals, Fact and Penalty, etc. In which, the penalty can be divided into three categories: Death-penalty, Life-imprisonment, and Imprisonment. Based on the original test data, this method combines the test data augmentation method with the multi-objective genetic algorithm to obtain the augmented test data of the penalty prediction model. The model framework is shown in Figure 1. The first part of the augmentation method: based on the text data describing the facts of the case, on the principle of not changing the meaning of the text, the original data is transformed by scrambling, inserting, and deleting the sentences in the text, and a large number of augmented data with the same label are obtained. The second part is the GA: first, the initial population is the augment data, and the individual is coded by decimal coding. The first-generation population is obtained by the selection, crossover, and mutation operation of GA. Then, starting from the second generation, the parent population and the offspring population are combined, and the precision, recall, F1 score, and importance of the test data are taken as the optimization objectives to evaluate the individuals. The hierarchy is divided according to the dominant relationship between individuals, and the crowding degree of individuals at the same level is calculated. Then, the superior individuals are selected to form a new parent population. Finally, the next-generation population is obtained by selection, crossover, and mutation operations of GA, and the above process will be repeated until the maximum number of iterations is reached.

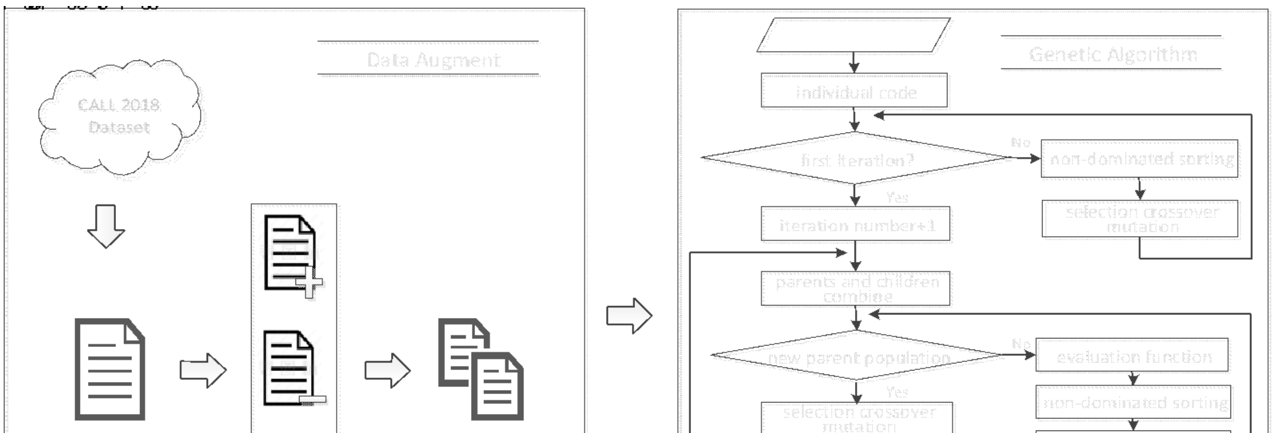

Figure 1. Method framework

3.1. Test Data Augment Method

Word order in text data has a great influence on sentence meaning, for example, "I love you" and "you love me" express different meanings. Therefore, in order not to change the meaning of the original text, this paper makes data augmentation for the sentences in the text. A large number of augmented data are generated by randomly scrambling the sentence order in the text, randomly deleting the sentence in the text, and randomly inserting the sentence in other text with the same label. The original data set is $D_0 = \{d_{01}, d_{02}, \cdots, d_{0l}\}$.

(1) Scrambling: The basic unit of this method is the complete sentence in the text. Since the sentence order has little influence on the text meaning of describing facts, randomly scrambling the sentence order in the original text can get the augmented text data with the same label and scale as the original data set, and then we get the scrambling data set: $D_1 = \{d_{11}, d_{12}, \cdots, d_{1l}\}$.

(2) Deletion: Since there are many redundant sentences in the text describing facts, these sentences have little influence on the meaning of the text, and deleting them will not affect the understanding of the case. Therefore, a sentence in the original text is deleted randomly. If the original text contains only one sentence, no processing is performed. By performing the same deletion operation for each text, the augmented text data with the same label and scale as the original data can be obtained, and then we get the deletion data set: $D_2 = \{d_{21}, d_{22}, \cdots, d_{2l}\}$.

(3) Insertion: Since there are many similar sentences in the text of the case with the same charge, the text data with the same charge label is divided into one category. Select a sentence from another text with the same label, and then insert it into the original data randomly. The augmented text data with the same label and scale as the original data can be obtained, and then we get the insertion data set: $D_3 = \{d_{31}, d_{32}, \cdots, d_{3l}\}$.

Based on the above analysis, we designed three kinds of data augmentation algorithms: scrambling, deletion, and insertion, as shown in algorithm 1.

Data augmentation is a very effective way to increase the size of data sets. Three new data sets with the same scale as the original data set can be obtained by the above three data augmentation methods. It is mixed with the original data set to obtain the augmented data set: $D = D_0 \cup D_1 \cup D_2 \cup D_3$.

3.2. Multi-Objective Genetic Algorithm

The multi-objective genetic algorithm takes all individuals in the population as the object and uses the multi-objective optimization technology to guide the encoded search space for an efficient search. Among them, the optimization objective and operation operators are the key technologies to realize the multi-objective genetic algorithm.

Randomly select the test data from the augmented data set to construct the initial population with the size of *n*. It is expressed as $x = \{x_1, x_2, \cdots, x_i, \cdots, x_n\}$. In which, the ith individual x is expressed as $x_i = \{x_{i,1}, x_{i,2}, \cdots, x_{i,j}, \cdots, x_{i,m}\}$, $x_{i,j}$ is the j^{th} test data of x_i, and m represents the size of x_i.

3.2.1. Optimization Objectives

The quality of test data has an important impact on the performance of the penalty prediction model. The most commonly used test data evaluation metrics are precision, recall, and F1 score [22]. Therefore, this paper uses these three evaluation metrics as the optimization target of GA. In addition, in order to improve the diversity of the test set, this paper gives the concept of individual importance from the distribution of test data and also takes it as the optimization objective.

(1) Precision: Among the samples predicted as correct by the model, the probability of actually correct samples is expressed as

$$Precision(x_i) = \frac{TP(x_i)}{TP(x_i) + FP(x_i)} \tag{1}$$

(2) Recall: In the actual correct sample, the probability predicted as the correct sample by the model is expressed as

$$Recall(x_i) = \frac{TP(x_i)}{TP(x_i) + FN(x_i)} \tag{2}$$

Algorithm 1 Data augmentation algorithm

Input:

Original text data set: D_0;

Output:

output scrambling text data set: D_1;

output delete text data set: D_2;

output insert text data set: D_3;

Augmented text data set: $D = D_0 \cup D_1 \cup D_2 \cup D_3$.

//Phase 1: Scrambling

for all d_t D_0 **do**

d_t. select sentence(i, j); $\Leftarrow$ randomly select sentence i and sentence j from d_t;

d_t. scrambling sentence(i, j); $\Leftarrow$ scrambling sentence i and sentence j in d_t;

end for

output scrambling text data set: D_1;

//Phase 2: Deletion

for all d_t D_0 **do**

d_t. select sentence(k); $\Leftarrow$ randomly select sentence k from d_t;

if $d_t - \{ k \}$) **then**

d_t. delete sentence(k); $\Leftarrow$ delete sentence k from d_t;

end if

end for

output delete text data set: D_2;

//Phase 3: insertion

for all d_t D_0 **do**

for d_s D_0 **do**

if d_s d_t **then**

d_s. select sentence(k); $\Leftarrow$ randomly select sentence k from d_s;

d_t. insert sentence(k); $\Leftarrow$ insert sentence k into d_t;

end if

end for

output insert text data set: D_3;

---

(3) F1 score: Comprehensively evaluate the precision and recall to make them reach the highest level at the same time. The obtained balance point is expressed as

$$F1(x_i) = \frac{2 \times Precision(x_i) \times Recall(x_i)}{Precision(x_i) + Recall(x_i)} \tag{3}$$

At the micro-level, $TP(x_i) = \sum_{j=1}^{m} TP(x_{i,j})$, which represents that the model correctly predicts the true sample x_i; $FP(x_i) = \sum_{j=1}^{m} FP(x_{i,j})$, which represents that the model correctly predicts false samples x_i; $FN(x_i) = \sum_{j=1}^{m} FN(x_{i,j})$, which represents that the model mispredicts the true sample x_i.

(4) Importance: This paper presents a statistical method of individual importance to evaluate the quality of individuals in the population. Main idea: the penalty as an important characteristic data to test the performance of the penalty prediction model. It can be divided into three categories: death-penalty, life-imprisonment, and imprisonment. If a certain kind of data appears frequently in individuals and rarely in populations, it is considered that this kind of data has good ability of classification, and the importance of the individuals is high. The importance of a certain type of data will increase proportionally with the frequency of its occurrence in the individual, and decrease inversely with the frequency of its occurrence in the population, which can effectively reflect the distribution of feature data and the importance of the individual. Therefore, based on the frequency of three types of data in individuals and populations, this paper comprehensively evaluates the importance of the individuals in which they are located and gives the expression of the importance of the individual is

$$Important(x_i) = \frac{3 \times \prod_{k=1}^{3} FD_k(x_i)}{\sum_{k=1}^{3} FD_k(x_i)} \tag{4}$$

$$FD_k(x_i) = \frac{1}{m} n_k(x_i) \times log \frac{|x|}{\sum_{i=1}^{n} n_k(x_i) + 1}, k = 1,2,3 \tag{5}$$

The total number of test data included by x_i is *m*. When k = 1,2,3, $n_k(x_i) = \sum_{j=1}^{m} n_k(x_{i,j})$, which refers to the number of times that three types of feature data of death penalty, life imprisonment, and imprisonment appear in x_i, and $FD_k(x_i)$ respectively shows their distribution.

3.2.2. Operators

The operators of the genetic algorithm include selection, crossover and mutation.

(1) Selection: The tournament selection strategy is used to randomly select *n*/2 individuals from the population n each time, and then the non-dominated sorting algorithm is used to obtain the Pareto optimal solution, from which the optimal individuals are selected to form the offspring population. Repeat until the new population size reaches n.

(2) Crossover: The circular cross method is adopted, and the cross process is shown in Figure 2. The first step is to randomly select a gene on the parent 1, find the gene number on the corresponding position of the parent 2, then return to the parent 1 to find the gene position with the same number. Repeat the previous work until a ring is formed, and the position of all genes in the ring is the last selected position; the second step is to generate a child with the gene selected by the parent 1, and ensure the corresponding position; the third step is to put the remaining genes into the offspring population.

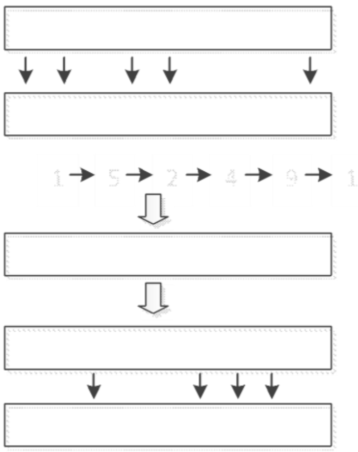

Figure 2. Cyclic cross process

(3) Mutation: Using the method of sequence number variation, we randomly select a gene position of the parent individual, delete the test case of this point, and then randomly insert a test case that does not duplicate the existing genes in the current individual to form a new offspring individual.

4. Experiment

4.1. Data Set

This paper uses China's AI and legal challenge data set (CALL), issued by the Supreme People's Court of China, with a total data set of 188,000. In which, the test data set is 33,000. The judgment document has a good structure, including the case information including the charge, date, address, criminals, fact, and penalty. In which, the penalty can be divided into three categories: death penalty, life imprisonment, and imprisonment.

4.2. Data Preprocessing

Since the original data is in the form of json, we extract the 'fact' and 'charge' fields according to the labels, serialize the 'charge' field and perform word segmentation on fact description text. We use Jieba as the word segmentation tool and remove meaningless stopping words in the text according to the commonly used stopping word list to form a new test data set: $D_0 = \{d_1, d_2, \cdots, d_l\}$.

4.3. Data Augmentation

The task of the model is to predict the penalty result according to the fact of a legal document. In order not to change the meaning of the original text, for the case with the same 'charge' label, we use three ways to augment test data in D_0 as follows:

(1) Scrambling: Randomly scrambling the sentence order in the original text can get the augmented text data with the same label, which is the scrambling data set: $D_1 = \{d_{11}, d_{12}, \cdots, d_{1l}\}$.

(2) Deletion: A sentence in the original text is deleted randomly. By performing the same deletion operation for each text, the augmented text data with the same label as the original data can be obtained, which is the deletion data set: $D_2 = \{d_{21}, d_{22}, \cdots, d_{2l}\}$.

(3) Insertion: Select a sentence from another text with the same label, and then insert it into the original data randomly. The augmented text data with the same label as the original data can be obtained, which is the insertion data set: $D_3 = \{d_{31}, d_{32}, \cdots, d_{3l}\}$.

Three new datasets with the same scale as D_0 can be obtained by using the above three methods to augment the data. Mix it with the original data set to get the augment data set: $D = D_0 \cup D_1 \cup D_2 \cup D_3$.

4.4. Data Search

Randomly select the test data from the augmented data set to construct the initial population with the size of n. It is expressed as $x = \{x_1, x_2, \cdots, x_i, \cdots, x_n\}$. In which, the ith individual x is expressed as $x_i = \left\{x_{i,1}, x_{i,2}, \cdots, x_{i,j}, \cdots, x_{i,m}\right\}$, $x_{i,j}$ is the jth test data of x_i, and m represents the number of test data x_i contained. Taking the precision, recall, F1 score, and importance of test data as the optimization objective, the multi-objective genetic algorithm is used to search for high-quality test data and obtain the augment test data set: $D' = \{d_1{}', d_2{}', \cdots, d_m{}'\}$.

4.5. Data Evaluation

The augment data obtained by this method and the random method are respectively injected into the penalty prediction model proposed by reference [4] for testing to evaluate the quality of augmented test data of the two methods with accuracy.

Accuracy is a general index to evaluate the performance of the deep learning model. It refers to the ratio of the number of samples correctly classified and the total number of samples for a given test data set. The expression is as follows

$$Accuracy(x_i) = \frac{TP(x_i) + TN(x_i)}{TP(x_i) + FP(x_i) + TN(x_i) + FN(x_i)} \tag{6}$$

At the micro-level, $TP(x_i) = \sum_{j=1}^{m} TP(x_{i,j})$, which represents that the model correctly predicts the true sample x_i; $FP(x_i) = \sum_{j=1}^{m} FP(x_{i,j})$, which represents that the model correctly predicts false samples x_i; $FN(x_i) = \sum_{j=1}^{m} FN(x_{i,j})$, which represents that the model mispredicts the true sample x_i; $TN(x_i) = \sum_{j=1}^{m} TN(x_{i,j})$, which represents that the model mispredicts the false samples x_i.

In this paper, GA is used to search *m* data from *D* as the augmented test data set $D' = \{d_1{}', d_2{}', \cdots, d_m{}'\}$; The random method selects *m* data from *D* randomly as the augmented test data set: $D'' = \{d_1{}'', d_2{}'', \cdots, d_m{}''\}$.

4.6. Experimental Results and Analysis

In order to observe the quality of different scales of augmented test data sets, we choose the number of test data as 10,000, 30,000, 60,000, 90,000 and 12,0000 respectively for experiments, and compare the quality of augmented test data of GA and the random method.

Table 1 shows the accuracy of the penalty prediction model tested by the augmented test data of the random method and GA. The results show that, compared with the random method, the accuracy is significantly reduced, which shows that the test data augmented by GA can test more defects of the model, reflecting the advantages of GA. With the increase of test data, the advantage to test the model accuracy of augmented data is first enhanced and then weakened. When the number of augmented test data reaches 60,000, the most obvious advantage is 3.5%. Compared with the random method, the comparative advantage of GA in testing the accuracy of the model through augmented data is shown in Figure 3. What we need to pay attention to is that with the expansion of test set scale, the quality of the test data increased by GA gradually decreases. When all the augmented test data are included, the ability of the two methods to test the accuracy of the model is basically the same.

Table 1. Accuracy

	Random Method	**Genetic Algorithm**
10 000	76.61%	74.08%
30 000	73.20%	70.45%
60 000	68.86%	65.36%
90 000	65.91%	64.38%
120 000	63.31%	63.17%

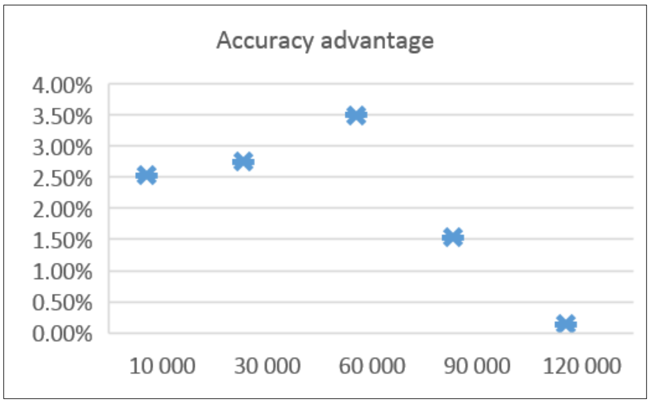

Figure 3. Accuracy advantage

5. Conclusion and Future Work

In this paper, GA is integrated into the research of judicial test data augmentation, which provides high-quality augmented test data for improving the reliability of the penalty prediction model. The data augmentation method based on GA has a more significant improvement in small-scale data sets, but the difference in large-scale data sets is not significant.

In future work and research, we will try more test data augmentation methods based on intelligent optimization, so as to obtain more diverse data, and carry out more experiments and research on the quality evaluation of augmented data. We hope that the augmented data like real data can fundamentally reduce the difficulty of data collection.

Acknowledgements

This work was supported by Mudanjiang Normal University Research Project (No.YB2020001, No.YB2020002, No.1354MSYTD005), Heilongjiang Provincial Department of Education Basic Research Business Expenses Project (No.1353MSYQN006).

References

1. S. Zhang, G. Yan, Y. Li, and J. Liu, "Evaluation of Judicial Imprisonment Term Prediction Model based on Text Mutation," in *Proceedings of 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion* (*QRS-C*), pp. 62-65, 2019
2. Y. Li, T. He, G. Yan, S. Zhang, and H. Wang, "Using Case Facts to Predict Penalty with Deep Learning," in *Proceedings of International Conference of Pioneering Computer Scientists, Engineers and Educators*, pp. 610-617, Springer, 2019
3. G. S. Becker, "Crime and Punishment: An Economic Approach," *The Journal Political Ewnomy*, Vol. 76, No. 2, pp. 169-217, 1968
4. T. He, H. Lian, Z. Qin, Z. Chen, and B. Luo, "PTM: A Topic Model for the Inferring of the Penalty," *Journal of Computer Science and Technology*, Vol. 33, No. 4, pp. 756-767, 2018
5. J. Liu, Z. Wang, G. Yan, and H. Lian, "Quality Measurement of Judgment Documents," in *Proceedings of 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion* (*QRS-C*), pp. 296-299, 2019

6. C. Lei, B. Hu, D. Wang, S. Zhang, and Z. Chen, “A Preliminary Study on Data Augmentation of Deep Learning for Image Classification,” in *Proceedings of the 11th AsiaPacific Symposium on Internetware*, pp. 1-6, 2019
7. G. Yan, Y. Li, S. Zhang, and Z. Chen, “Data Augmentation for Deep Learning of Judgment Documents,” in *Proceedings of International Conference on Intelligent Science and Big Data Engineering*, pp. 232-242, Springer, 2019
8. L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” arXiv Preprint arXiv:1712.04621, 2017
9. C. Xia, Y. Zhang, L. Wan, Y. Song, and N. Xiao, “Test Data Generation of Path Coverage based on Negative Selection Genetic Algorithm,” *Acta Electronica Sinica*, Vol. 47, No. 12, pp. 2630-2638, 2019
10. H. Wang, T. He, Z. Zou, S. Shen, and Y. Li, “Using Case Facts to Predict Accusation based on Deep Learning,” in *Proceedings of 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion* (*QRS-C*), pp. 133-137, 2019
11. C. -L. Liu and C. -D. Hsieh, “Exploring Phrase-based Classification of Judicial Documents for Criminal Charges in Chinese,” *Foundations of Intelligent Systems*, Vol. 4203, pp. 681-690, Springer, Heidelberg, 2006
12. B. Luo, Y. Feng, J. Xu, X. Zhang, and D. Zhao, “Learning to Predict Charges for Criminal Cases with Legal Basis,” arXiv Preprint arXiv:1707.09168, 2017
13. H. Zhong, G. Zhipeng, C. Tu, C. Xiao, Z. Liu, and M. Sun, “Legal Judgment Prediction via Topological Learning,” in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3540-3549, 2018
14. H. Inoue, “Data Augmentation by Pairing Samples for Images Classification,” arXiv Preprint arXiv:1801.02929, 2018
15. R. Takahashi, T. Matsubara, and K. Uehara, “Data Augmentation using Random Image Cropping and Patching for Deep CNNs,” *IEEE Transactions on Circuits and Systems for Video Technology*, 2019
16. Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random Erasing Data Augmentation,” arXiv preprint arXiv:1708.04896, 2017
17. A. W. Yu, D. Dohan, M. Luong, R. Zhao, K. Chen, M. Norouzi, et al., “Qanet: Combining Local Convolution with Global Self-Attention for Reading Comprehension,” in *Proceedings of 6th International Conference on Learning Representations, ICLR 2018*, Vancouver, BC, Canada, April 30 - May 3 2018
18. W. Y. Wang and D. Yang, “That's So Annoying!!!: A Lexical and Frame-Semantic Embedding based Data Augmentation Approach to Automatic Categorization of Annoying Behaviors using #Petpeeve Tweets,” in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics*, pp. 2557-2563, Lisbon, Portugal, 2015
19. J. Wang and L. Perez, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” *Convolutional Neural Networks Vis. Recognit*, Vol. 11, 2017
20. K. Deb, A. Pratap, S. Agarwal, and T. Meyadvan, “A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-up,” *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, pp. 182-197, 2002
21. B. T. Ma, X. Li, Y. Xia, and Y. N. Zhang, “Autonomous Deep Learning: A Genetic DCNN Designer for Image Classification,” *Neurocomputing*, Vol. 379, pp. 152-161, 2020
22. G. Yan, Y. Li, S. Zhang, and S. Shen, “Law Article Prediction based on Deep Learning,” in *Proceedings of International Conference on Software Quality, Reliability and Security Companion* (*QRS-C*), pp. 281-284, 2019

Chunyan Xia is an associate professor of Department of Computer and Information Technology at Mudanjiang Normal University, her research interests include search-based software engineering, information processing and data mining.

Xingya Wang is an associate professor of College of Computer Science and Technology at Nanjing Tech University, his research interests include blockchain analysis and test, software defect location.

Yan Zhang is a professor of Department of Computer and Information Technology at Mudanjiang Normal University, her research interests include search-based software engineering.

Hao Yang is an assistant of Department of Computer and Information Technology at Mudanjiang Normal University, his research interests include search-based software engineering.